High-resolution image reconstruction with latent diffusion models from human brain activity

Abstract

Reconstructing visual experiences from human brain activity offers a unique way to understand how the brain represents the world, and to interpret the connection between computer vision models and our visual system. While deep generative models have recently been employed for this task, reconstructing realistic images with high semantic fidelity is still a challenging problem. Here, we propose a new method based on a diffusion model (DM) to reconstruct images from human brain activity obtained via functional magnetic resonance imaging (fMRI). More specifically, we rely on a latent diffusion model (LDM) termed Stable Diffusion. This model reduces the computational cost of DMs, while preserving their high generative performance. We also characterize the inner mechanisms of the LDM by studying how its different components (such as the latent vector of image Z, conditioning inputs C, and different elements of the denoising U-Net) relate to distinct brain functions. We show that our proposed method can reconstruct high-resolution images with high fidelity in straight-forward fashion, without the need for any additional training and fine-tuning of complex deep-learning models. We also provide a quantitative interpretation of different LDM components from a neuroscientific perspective. Overall, our study proposes a promising method for reconstructing images from human brain activity, and provides a new framework for understanding DMs. Please check out our webpage at https://sites.google.com/view/stablediffusion-with-brain/.